System Identification, Lecture 3

Kristiaan Pelckmans (IT/UU, 2338)

Course code: 1RT875, Report code: 61806,
F, FRI Uppsala University, Information Technology

22 January 2010

Lecture 3

- Nonparametric Methods (Ch. 3)
- Input Signals (Ch. 4)
- Model Parametrizations (Ch. 5)

System Identification

'Obtain a model of a system from measured inputs and outputs'

Type of model depends on purpose, application and system. Often we can assume that the *true* system can be described as a LTI system.

$$y(t) = G_0(q^{-1})u(t) + v(t),$$

or equivalently

$$y(t) = \sum_{\tau=1}^{\infty} g_0(\tau)u(t-\tau) + v(t)$$

Q: How to approximate $G_0(q^{-1})$ from measurements?

Parametric Models

Postulate a model class *parametrized* by $\theta \in \Theta$:

$$\mathcal{M}_\Theta = \left\{ G(q^{-1}, \theta) : \theta \in \Theta \right\}$$

- Easy to use for simulation, control design, etc. ...
- Often accurate models.
- Ex. FIR model

$$y(t) = u(t) + b_1 u(t-1) + \cdots + b_\tau u(t-\tau)$$

or $y(t) = G_F(q^{-1}, \theta)$ with

$$G_F(q^{-1}, \theta) = 1 + b_1 q^{-1} + \cdots + b_\tau q^{-\tau}, \theta = (b_0, \ldots, b_\tau)^T \in \mathbb{R}^{\tau+1}$$

Q.: Can we determine Q_0 without postulating a model?

Nonparametric Identification

Nonparametric models: Determine G_0 without postulating $\mathcal{M}_\Theta$.

- Simple to obtain
- Graphs, curves or tables, but often no simulation
- Often used to validate parametric models
- Transient, correlation, frequency and spectral analysis.

Transient Analysis

Impulse response analysis: Apply the input

$$u(t) = \begin{cases} k & t = 0 \\ 0 & \text{else} \end{cases}$$

to the system G_0. This gives the output signal

$$y(t) = kg_0(t) + v(t)$$

and this motivates the impulse estimate for all $\tau \geq 0$

$$\hat{g}(\tau) = \frac{y(\tau)}{k}.$$

Transient Analysis (Ct'd)

Step response analysis: Apply the input

$$u(t) = \begin{cases} k & t \geq 0 \\ 0 & \text{else} \end{cases}$$

to the system G_0. This gives the output signal

$$y(t) = k \sum_{k=1}^{t} g_0(k) + v(t)$$

and this motivates the impulse estimate for all $\tau \geq 1$

$$\hat{g}(\tau) = \frac{y(\tau) - y(\tau - 1)}{k}.$$

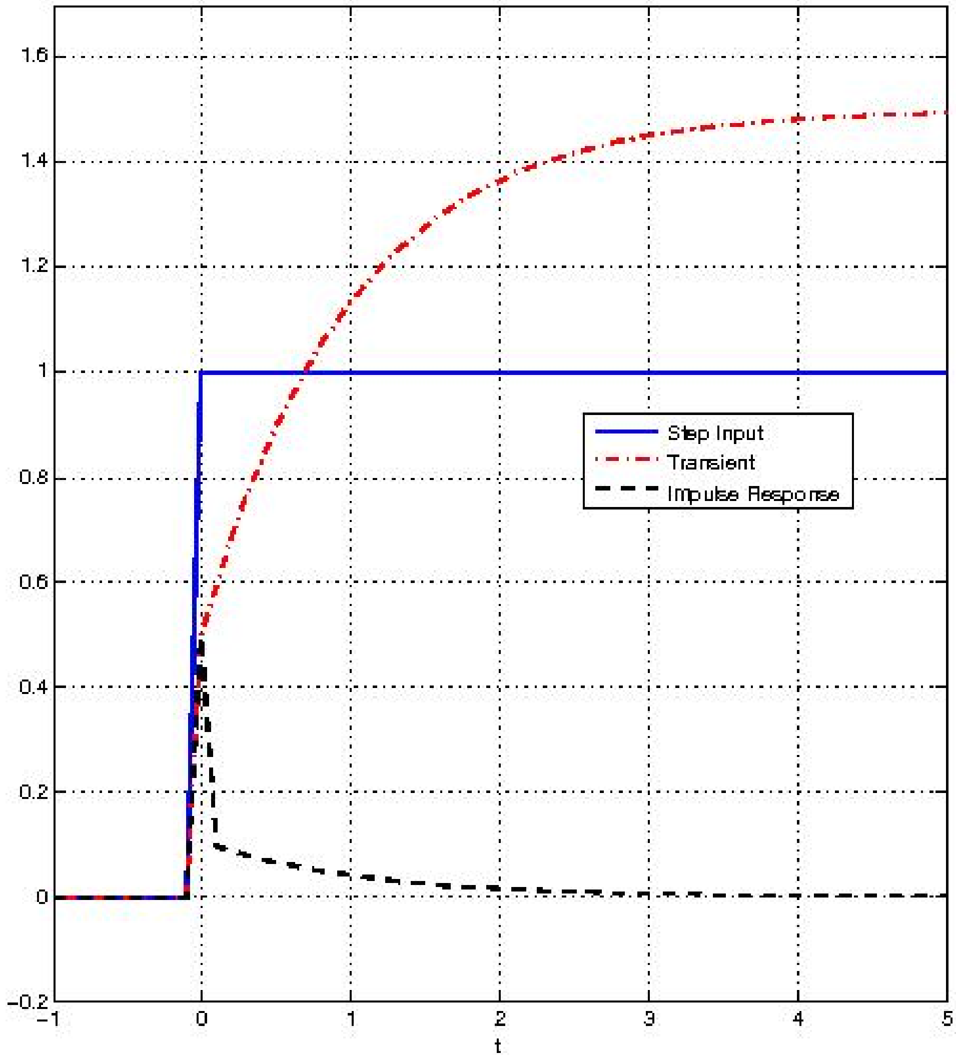

Figure 1: Transient Behavior of G_0 on a step input $u(t)$

Transient Analysis

- Input taken as impulse or step.
- 'Model' consists of recorded outputs $y(t)$, or estimates of $g_0(t)$
- Convenient to derive crude models. Gives estimates of time-constants time-delays and static gain.
- Sensitive to noise.
- Poor excitation.

Correlation Analysis

System

$$y(t) = \sum_{k=1}^{\infty} g_0(k) u(t-k) + v(t)$$

where $u(t)$ is a stochastic process independent of $v(t)$. Multiplication with $u(t')$ of both sides and taking expectations gives ($\tau = 0, \ldots, t$) that

$$r_{uy}(\tau) = \sum_{k=1}^{\infty} g_0(k) r_{uu}(\tau - k)$$

which is known as the *Wiener-Hopf* equation.

In practice, truncate the sum and solve the linear systems of equations

$$\hat{r}_{uy}(\tau) = \sum_{k=1}^{M} \hat{g}_c(k) \hat{r}_{uu}(\tau - k)$$

Estimates of the covariance functions $\hat{r}_{uy}$ and $\hat{r}_{uu}$ gives (for $\tau \geq 0$)

- First choice

$$\hat{r}_{uy}(\tau) = \frac{1}{N} \sum_{k=1}^{N-\tau} y(k+\tau)u(k).$$

- Second choice

$$\hat{r}_{uy}(\tau) = \frac{1}{N-\tau} \sum_{k=1}^{N-\tau} y(k+\tau)u(k).$$

Which one to prefer?

Frequency Analysis

Estimate $G_0(e^{i\omega})$. Apply input signal

$$u(t) = \alpha \cos(\omega t)$$

to $G_0(e^{i\omega})$. This yields output signal

$$y(t) = \alpha \left|G_0(e^{i\omega})\right| \cos(\omega t + \varphi) + v(t)$$

- Repeat experiment for different frequencies ω $(t = 1, \ldots, N)$
- Determine the phase shift φ and the amplitude of the output.
- Results in a Bode plot $\left\{\left|G_0(e^{i\omega})\right|\right\}_\omega$ and $\left\{\angle G_0(e^{i\omega})\right\}_\omega$
- Sensitive to noise. requires long experiments.
- Gives basic information about the system.

Spectral Analysis

- The Wiener-Hopf equation in the frequency domain is given as

$$\phi_{uy}(\omega) = G(e^{-i\omega})\phi_u(\omega)$$

- An estimate of the transfer function can be given as

$$\hat{G}(e^{-i\omega}) = \frac{\phi_u(\omega)}{\phi_{uy}(\omega)}$$

- Use estimate of the spectral densities, e.g.

$$\hat{\phi}(\omega) = \frac{1}{2\pi N} \sum_{\tau=-N}^{N} \hat{r}_{yu}(\tau) e^{-i\tau\omega}$$

- Errors in $\hat{r}_{uy}$ contaminate $\rightarrow$ not consistent!
  - N large, then total norm of error is large even if $\hat{r}_{uy}$ is small for all τ.
  - $\hat{r}_{uy}$ decreases slowly, then poor estimates of $\hat{r}_{uy}$ for large τ.

- Better estimates obtained if 'window $w(\tau)$' used

$$\hat{\phi}(\omega) = \frac{1}{2\pi N} \sum_{\tau=-N}^{N} \hat{r}_{yu}(\tau) w(\tau) e^{-i\tau\omega}$$

- Choice of window is a trade-off between bias and variance (high resolution and reducing erratic fluctuations)

Summary - Nonparametric Methods

- Results often in graph or table (step response, transfer function, ...)
- Transient analysis (step- and impulse response)
- Frequency analysis (sinusoidal input)
- Correlation analysis
- Spectral analysis (transfer function)
- Useful for obtaining crude estimates of time-constants, cut-off frequencies etc. for model validation.

Input Signals (Ch. 5)

The quality of the estimated model depends on the choice of input signal.

Examples:

- Step function
- Pseudo-random binary sequences (PRBS)
- Autoregressive moving average process (ARMA)
- Sum of sinusoids.

Most often the input signal is characterized by its first and second moments:

$$\begin{cases} m = E[u(t)] \\ r(\tau) = E\left[(u(t) - m)(u(t) - m)^T\right] \end{cases}$$

and/or its spectral density:

$$\phi(\omega) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} r(\tau) e^{-i\tau\omega}$$

Rem. for stationary signals

$$m = \frac{1}{N} \sum_{t=1}^{N} u(t)$$

Step function

$$u(t) = \begin{cases} k & t = 0 \\ 0 & \text{else} \end{cases}$$

Properties

- Mostly used for transient analysis: overshoot, static gain, major time-constants.
- Limited use for parametric modeling.

Pseudo-Random Binary Sequences (PRBS)

A PRBS $(u(t))_t$ is a periodic, deterministic signal with white noise-like properties.

$$u(t) = \operatorname{rem}\left(A(q^{-1})e(t), 2\right)$$

Properties

- The signal takes values $\{0, 1\}$ in a fashion dictated by A.
- Spectral properties are determined by $A(q)$ and in particular by the period length $M = 2^n - 1$.
- Deterministic sequence behaving as noise (reproducible).

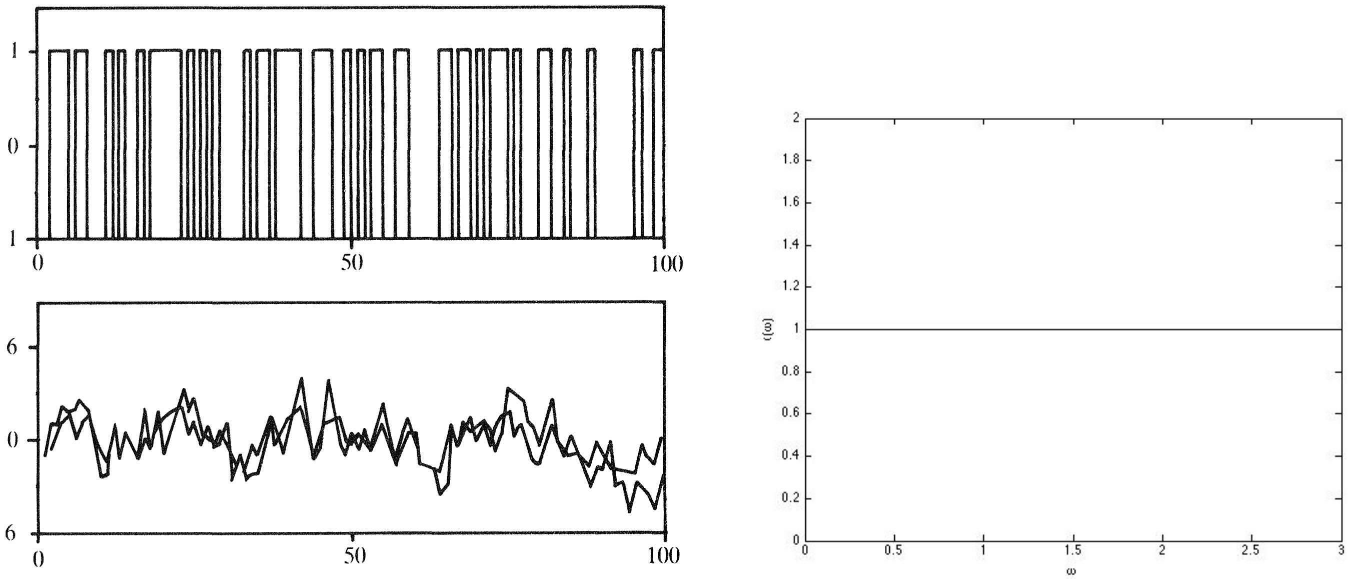

Figure 2: PRBS signal taking values in $\{-1, 1\}$, $M = \infty$. Realization (left), Spectral density (right).

ARMA Process

$$A(q^{-1})y(t) = C(q^{-1})e(t)$$

where $e(t)$ is white noise with $E[e(t)] = 0$ and $E[e(t)e(s)] = \lambda^2\delta_{ts}$.

Properties

- The signal $u(t)$ can be obtained by filtering $e(t)$.
- The filters (A, C) can be tuned to possess (almost) any frequency characteristics.
- The spectral density of an ARMA process $y(t)$ is given as

$$\phi_y(\omega) = \frac{\lambda^2}{2\pi}\left|\frac{C(e^{i\omega})}{A(e^{i\omega}}\right|^2$$

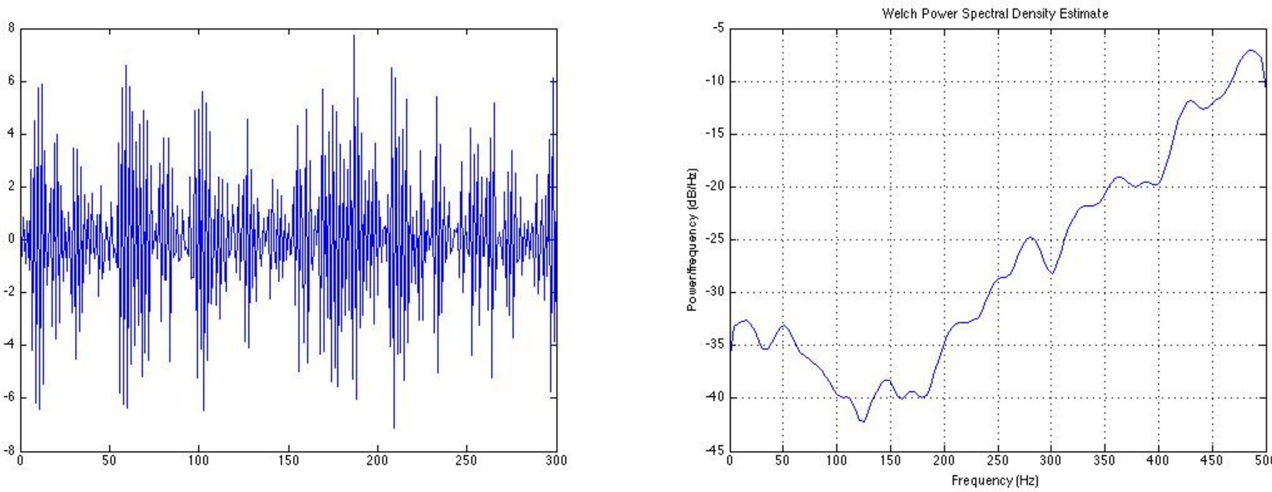

Figure 3: ARMA process. Realization (left), Spectral density (right).

Sum of Sinusoids

$$u(t) = \sum_{m=1}^{M} a_m \sin(\omega_m t + \varphi_m).$$

Properties

- User parameters a_m, ω_m, φ_m.
- Covariance function given as

$$r(\tau) = \sum_{m=1}^{M} \frac{a_m^2}{2} \cos(\omega_m t + \varphi_m).$$

- Spectral Density function given as

$$\phi(\omega) = \sum_{m=1}^{M} \frac{a_m^2}{2} \left[\delta(\omega - \omega_m) + \delta(\omega - \omega_m)\right].$$

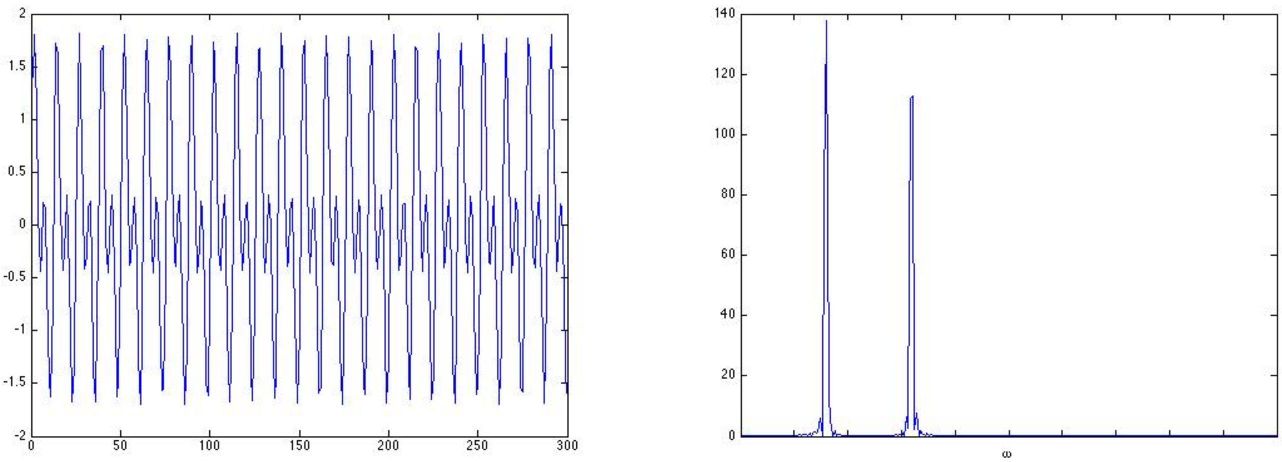

Figure 4: Sum of 2 sinusoids. Realization (left), Spectral density (right).

Persistent Excitation

In order to obtain a good estimate of a (parametric) model, the input signal has to be 'rich' enough so that all 'modes' of the system are excited.

An inout is said to be persistently exciting (PE) if:

- The following limit exists for all τ

$$r_u(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N-\tau} u(t+\tau)u^T(t)$$

Rem. $u(t)$ ergodic implies that for any t

$$r_u(\tau) = E[u(t+\tau)u^T(t)]$$

- The matrix $\mathbf{R}_u(n)$

$$\mathbf{R}_u = \begin{bmatrix} r_u(0) & r_u(1) & \dots & r_u(n-1) \\ r_u(1) & r_u(0) & \dots & \vdots \\ \vdots & & \ddots & \\ r_u(n-1) & \dots & & r_u(0) \end{bmatrix}$$

is positive (strictly) definite.

- Or, $\det(\mathbf{R}_u(n)) \neq 0$.

- Or $u(t)$ is PE of order n if $\phi_u(\omega) \neq 0$ on at least n points on the interval $-\pi < \omega < \pi$.

An input signal is PE of order $2n$ can be used to consistently estimate parameters of a model of order $\leq n$.

- A step function that is PE of order 1
- A PRBS with period M is PE of order M.
- An ARMA process is PE of any finite order.
- A sum of m sinusoids is PE of order $2M$ (if $\omega_m \neq 0, -\pi, \pi$)

Another important observation!

A parametric model becomes more accurate in the frequency region where the input signal has a major part of its energy.

A physical process is often of low frequency character $\rightarrow$ use low-pass filtered signal as input.

Summary - Input Signals

- The choice of input signals determines the quality of the estimate.
- The estimated model is more accurate in frequency regions where the input signal contains much energy.
- An input signal has to be 'rich' enough to excite all interesting modes of the system (PE of sufficiently high order).
- In practice there might be restrictions on the input.

Model Parametrization (Ch. 6)

Mathematical models can be derived from:

- Physical models
- Identification

Classification of mathematical models:

- SISO - MIMO.
- Linear - Nonlinear models.
- Parametric - Nonparametric.
- Time-invariant - time-varying.
- Time-domain - Frequency domain.
- Discrete-Time - Continuous-Time.
- Deterministic - Stochastic.

General Model Structure (SISO)

$$y(t) = G\left(q^{-1}, \theta\right) u(t) + H\left(q^{-1}, \theta\right) e(t)$$

- where

$$G\left(q^{-1}, \theta\right) = \frac{A\left(q^{-1}\right)}{B\left(q^{-1}\right)} = \frac{b_1 q^{-n_k} + b_2 q^{-n_k - 1} + \cdots + b_{n_b} q^{-n_k - n_b}}{1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}}$$

- and

$$H\left(q^{-1}, \theta\right) = \frac{C\left(q^{-1}\right)}{D\left(q^{-1}\right)} = \frac{1 + c_1 q^{-1} + \cdots + c_{n_c} q^{-n_c}}{1 + d_1 q^{-1} + \cdots + d_{n_d} q^{-n_d}}$$

- and $e(t)$ is white noise with variance λ^2 and

$$\theta = \left(a_1, \ldots, a_{n_a}, b_1, \ldots, b_{n_b}, c_1, \ldots, c_{n_c}, d_1, \ldots, d_{n_d},\right)^T$$

- Often $\lambda^2 = \lambda^2(\theta)$.

Assumptions

- Time delay $n_k \geq 1 \rightarrow G(0,\theta) = 0$ (often also $G(0,\theta) = 0$).
- $G^{-1}(q^{-1},\theta)$ and $H^{-1}(q^{-1},\theta)$ are asymptotically stable (...). Often also $H(q^{-1},\theta)$ needs to be asymptotically stable.

General Model Structures (Ct'd)

Commonly used simplified models

- ARMAX

$$A(q^{-1})y(t) = B(q^{-1})u(t) + C(q^{-1})e(t).$$

Here $A(q^{-1})$ describes the dynamics. Both inputs and noise are governed by the same dynamics.

- ARX

$$A(q^{-1})y(t) = B(q^{-1})u(t) + e(t).$$

- FIR

$$y(t) = B(q^{-1})u(t) + e(t).$$

- OE

$$y(t) = \frac{B(q^{-1})}{A(q^{-1})}u(t) + e(t).$$

General Model Structures (Ct'd)

Time series models (no 'input' signal $u(t)$)

- ARMA
$$A(q^{-1})y(t) = C(q^{-1})e(t).$$
- AR
$$A(q^{-1})y(t) = e(t).$$
- MA
$$y(t) = C(q^{-1})e(t).$$

Time series models are useful in various disciplines, e.g. economy, astrophysics, speech, etc... .

Uniqueness and Identifiability

Uniqueness: Let the *true* system $\mathcal{S}$ be described by G_0, H_0 and λ_0^2.

Introduce the set

$$\mathcal{D}_T = \left\{\theta \;\middle|\; G_0 = G(q^{-1},\theta), H_0 = H(q^{-1},\theta), \lambda_0^2 = \lambda^2(\theta)\right\}$$

- $|\mathcal{D}_T| = 0$ underparametrized model structure
- $|\mathcal{D}_T| > 1$ overparametrized model structure (numerical problems are likely to occur)
- $|\mathcal{D}_T| = 1$ Ideal case. The system has a unique description as θ_0

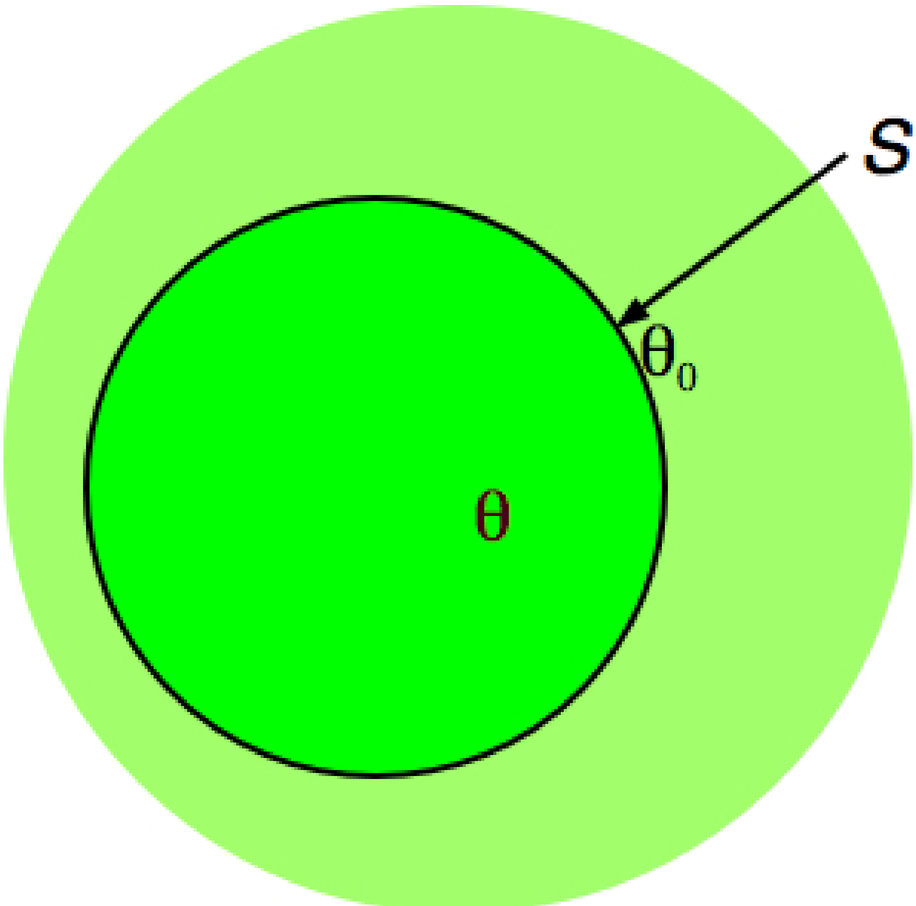

Figure 5: Model structure (Green area), actual 'true' system $\mathcal{S}$, estimate θ and best approximation θ_0.

Uniqueness and Identifiability (Ct'd)

Identifiability:

- System Identifiability (SI): $|\mathcal{D}_T| > 0$, and $\hat{\theta} \in \mathcal{D}_T$ if $N \to \infty$.
- Parameter Identifiability (PI): If the system is SI and $|\mathcal{D}_T| = 1$ (or $\hat{\theta} \to \theta_0$).

In other words, if the choice of *model, input signal and identification method* makes the estimated parameter vector $\hat{\theta}$ converge (with probability 1 as $N \to \infty$) to a parameter vector that perfectly describes the system as the number of datapoints tends to infinity, then the system is **System Identifiability (SI)**. If the system is uniquely described by an element in the model structure *and* is SI then the system is said to be **parameter identifiable (PI).**

Summary - Model Parametrizations

- It is essential that the model structure suits the actual system.
- Many standard model structures are available, each one using a different approach of modeling the influence of input $u(t)$ and disturbance signals $e(t)$.
- Finding the correct, or the best, model structure $\mathcal{M}$ and model order(s) $(n_a, n_b, n_c, n_d)^T$ is normally an iterative procedure (see Ch. 11)
- A model should ideally be unique and the complete experimental setup should be such that the system is PI.
- Not included: Ex. 6.3, 6.4, 6.6, continuous-time models. "Kursivt" ex. 6.5.